Stochastic Strategy Adjustment in Coordination Games

Kosfeld, M.

Publication date:
1999

Citation for published version (APA):
Kosfeld, M. (1999). *Stochastic Strategy Adjustment in Coordination Games*. (FEW Research Memorandum; Vol. 775). Microeconomics.

Stochastic Strategy Adjustment in Coordination Games

Michael Kosfeld*

CentER & Department of Econometrics,
Tilburg University, P.O.Box 90153, 5000 LE Tilburg, The Netherlands
(Email: M.Kosfeld@kub.nl)

Abstract

We explore a model of equilibrium selection in coordination games, where agents stochastically adjust their strategies to changes in their local environment. Instead of playing perturbed best–response, we assume that agents follow a rule of *"switching to better strategies more likely"*. We relate this behavior to work of Rosenthal (1989) and Schlag (1998). Our main results are that both strict Nash equilibria of the coordination game correspond to stationary distributions of the process, hence evolution of play is not ergodic, but instead depends on initial conditions. However, coordination on the risk–dominant equilibrium occurs with probability one whenever the initial share of agents playing the risk–dominant strategy has at least some positive measure, how ever small, within the whole population.

Journal of Economic Literature Classification: C72

Keywords: equilibrium selection, coordination game, evolution, strategy adjustment.

*This work has profited from several discussions with many friends and colleagues. In particular I would like to thank Hans Föllmer, Avner Shaked, Karl Schlag, and Mario Türnich for helpful comments on an earlier version of the paper. I am grateful to an anonymous referee for very useful suggestions. This version: November, 1998.

1. Introduction

The seminal work of Kandori, Mailath, and Rob (1993), henceforth denoted as KMR, and Young (1993) has attracted much interest in evolutionary models for equilibrium selection in coordination games. Subsequent models have refined this work by introducing local interaction (e.g., Ellison, 1993; Blume, 1993, 1995), or by enlarging the strategy space of an agent (Ely, 1995; Bhaskar and Vega-Redondo, 1997; Kim and Sobel, 1995). The present paper belongs to the first category in featuring local interaction. It follows a new line in studying alternative ideas for modelling the individual behavior of an agent.

The usual story in the evolutionary approach is that there is a large population of agents, each facing a situation of repeated interaction with other agents. The interaction is modelled as a symmetric 2×2 coordination game where agents are restricted to pure strategies. The evolution of play within the population is driven by the assumption that agents may switch strategies. Since opponents may change their strategy, too, each agent repeatedly plays the coordination game against a changing mixture of strategies. An agents task is to *adjust* his strategy to the environment he faces.

The original assumption of KMR (1993) and Young (1993) is that agents adjust their strategy by playing *perturbed best–response*. With high probability they play a best–response to their environment, with remaining low probability they simply play random. The first part is based on the idea that agents are influenced by payoff differences, the second part captures noisy behavior and models, e.g., individual mistakes or deliberate experimentation of an agent. Based on this assumption the surprisingly strong result is that evolution selects the risk–dominant equilibrium as defined by Harsanyi and Selten (1988).[1]

Perhaps the strongest objection to this result has been formulated by Bergin and Lipman (1996) who show that the equilibrium selection result is based on the specific assumption that random play is sufficiently similar in different states of the process. In fact they show that, if one allows the noise rate to depend on the state of the process then every invariant distribution of the noiseless process and thus every strict Nash equilibrium of the coordination game can be selected.[2]

One possible way to proceed is to make the noise part explicit by modelling the economic, social, or psychological source of it. A recent approach in this direction is van Damme and Weibull (1998), where agents rationally choose to make mistakes because it is too costly to avoid these mistakes completely. The resulting endogenous noise rates are shown to be sufficiently similar across states, which establishes the result of risk–dominance. An alternative is to leave the

[1]Papers in the second category mentioned above obtain the payoff–dominant equilibrium, which is a result that follows directly from the enlargement of the strategy space. See also our section 11.

[2]Blume (1994) argues in a similar way. He analyzes different versions of noise in order to find criteria for invariance of the selection result with respect to the noise process.

paradigm of perturbed best–response behavior and choose a different model for agents' adjustment, which nonetheless features important properties of the original approach. This is the way we proceed in this paper.

Intuitively speaking, pure best–response adjustment says: *adjust to best strategies with probability one*. This holds true even if the other strategy earns only infinitesimally larger payoffs. Hence adjustment is very payoff–sensitive and in fact deterministic. At any time t an agents probability to change his strategy is either 1 if the other strategy is a best–response, 0 if not. Our starting point is to study a smoothened version of best-response, which says: *adjust to better strategies more likely*. In a static framework this simple intuitive idea is related to a model of Rosenthal (1989).[3] From a decision theoretic point of view it resembles much the behavioral notion in *proportional imitation* as introduced by Schlag (1998), although we do not consider actual imitation in this model.

The first consequence of our assumption is that the strategy adjustment of an agent is less payoff–sensitive. Probabilities to switch lie within the whole interval $[0, 1]$ rather than in the subset $\{0, 1\}$. In effect, many times both strategies will have positive probability to be played. This corresponds to the randomness effect of noise in perturbed best–response. However, the second and main feature of our approach is that agents are still influenced by payoff differences, which we see as the essence of pure best–response behavior. While the latter assumes that infinitesimally small payoff differences are weighted in the same way as large payoff differences our assumptions say that payoff differences matter the more the larger they are.

Based on this approach our main results show that agents are more likely to coordinate on the risk–dominant equilibrium. In this sense our model supports the result of KMR (1993), Young (1993), and others and moderates the critique of Bergin and Lipman (1996). However, contrary to most other approaches, in our model the risk–dominated equilibrium still corresponds to a stationary distribution of the stochastic process. In other words, the process is not ergodic and, in fact, basins of attraction of both equilibria are non–empty. Thus, the question on which equilibrium agents will actually coordinate depends on initial conditions. Note that this is true under pure best–response adjustment as well. In order to obtain an equilibrium selection result noise needs to be introduced, which turns the process into an ergodic process and eliminates one of the two equilibria. This, however, produces a common paradox as it has been indicated, for instance, by Blume (1993, p415). While the theory stated that all players will *always* choose the risk–dominant strategy in the future, computer simulations showed

> "that the outcome depends strongly on the initial conditions of the process. If the initial frequency of down [i.e. the risk–dominated strategy] is sufficiently high, the process converges to all players choosing down."

A contribution of our approach is to resolves this paradox, by answering the equilibrium selection problem without necessarily eliminating one of the two equilibria.

[3]We are grateful to Bob Rosenthal for indicating us to this work.

The paper is organized as follows. In the following section we define our model of local interaction. Section 3 introduces the class of coordination games we want to look at. In section 4 the idea of stochastic strategy adjustment is made precise and our two main behavioral assumptions are formulated. In view of a better motivation these assumptions are related to other work in section 5. Section 6 then pins down the model. Results are obtained and discussed in sections 7 to 10. Section 11 concludes.

2. Local Interaction

Similarly to other models (Blume 1993, 1995; Ellison, 1993) we consider a spatial model of local interaction. Precisely, we assume an infinite population of agents that are located on the n-dimensional integer lattice $\mathbb{Z}^n$. The dimension of the lattice can have any value $n \geq 1$. Results in our model do not depend on n. By identifying each agent with his or her location the space $\mathbb{Z}^n$ represents the population of agents. Typically, agents will be denoted as $x, y, z \in \mathbb{Z}^n$.

Every agent is assumed to interact with a finite set of other agents, his so–called *neighbors*. For every agent $x \in \mathbb{Z}^n$ we define the neighborhood to be given by $N(x) := \{y \in \mathbb{Z}^n \mid |y \Leftrightarrow x| = 1\}$ where $|\cdot|$ denotes the Euclidean distance within $\mathbb{Z}^n$. Thus neighbors are agents that are one step away in at most one of n dimensions. Sometimes this kind of neighborhood interaction is also called *nearest neighbor interaction*. If n equals one, neighbors are located both to the right and to the left of an agent. For $n = 2$, the set of neighbors consists, in addition, of those agents that are located to the top and to the bottom of an agent.

Of course, there are various other possibilities for defining appropriate neighborhood structures, even if one sticks to the general assumption of an n-dimensional lattice. And intuition suggests that different structures will support different outcomes. Important research therefore focuses on robustness–checks with respect to different neighborhood structures. An early approach in this direction has been made by Ellison (1993) in comparing global with local interaction. A recent discussion of general neighborhood structures is given in Morris (1996) and Young (1998).

Yet, in this model we restrict our analysis to the nearest neighbor interaction as defined above. The reason for doing this is simply that this neighborhood structure has already been studied in other models, as well (Blume, 1993, 1995; Ellison, 1993). As we do not want to focus on the role of the neighborhood structure itself, but rather on the behavior of agents within such a neighborhood structure, this gives us the possibility to relate our findings to those of others without any confusion about possible differences in underlying neighborhood structures. Our main contribution shall be a check of robustness with respect to an agents behavior rather than with respect to the considered neighborhood structure.

3. Coordination Games

We consider the class of symmetric 2×2 coordination games that are given by the payoff matrix in Figure 1.

	Top	*Bottom*
Top	a, a	c, d
Bottom	d, c	b, b

Figure 1: A Coordination Game

We assume all values to be finite and both, $a > d$ and $b > c$, hence (Top, Top) and $(Bottom, Bottom)$ are the two strict Nash equilibria of the game. There exists another symmetric Nash equilibrium in mixed strategies, where both players put probability $\frac{b-c}{a-d+b-c}$ on strategy *Top*. However, we are not going to focus on this equilibrium since we restrict players to play pure strategies only.[4]

Time is modelled continuously. At any time $t \in \mathbb{R}_0^+$ to each agent $x \in \mathbb{Z}^n$ there is assigned one of the two possible actions, *Top*, henceforth denoted by T, or *Bottom*, henceforth denoted by B. The collection of actions over the whole population at time t is given by a mapping

$$\begin{aligned} \xi_t : \mathbb{Z}^n &\rightarrow \{T, B\} \\ x &\mapsto \xi_t(x), \end{aligned}$$

where $\xi_t(x)$ denotes the action of agent x at time t. A mapping ξ_t is also called a *configuration*. Denote X the set of all possible configurations.

Agents continuously and uniformly interact with their neighbors. We assume that at any time t each agent is sequentially matched with his $2n$ neighbors.[5] In each single match the coordination game is played with every agent choosing his assigned action. Denote for agent $x \in \mathbb{Z}^n$ by $\pi^s(x, \xi_t)$ the accumulated payoff from these matches, when the play of the population, in particular of his neighbors, is determined by configuration ξ_t and agent x plays strategy $s \in \{T, B\}$, hence $\xi_t(x) = s$. Thus,

$$\pi^s(x, \xi_t) = \sum_{y \in N(x)} G(s, \xi_t(y)), \tag{1}$$

where $G(\cdot, \cdot)$ calculates the payoff from the matrix in Figure 1.

[4] Bhaskar and Vega-Redondo (1997) distinguish between cases where $d \geq b$ and $d < b$. Games in the first subclass are called *stag-hunt games*, games in the second *pure coordination games*. Our results hold for both classes.

[5] In this paper we do not model the actual matching procedure but concentrate on accumulated payoffs.

4. Stochastic Strategy Adjustment

In our model we pursue the idea that agents stochastically adjust their strategy to a changing environment given by the play in their local neighborhood. We do not follow the usual approach in the evolutionary literature that studies boundedly rational behavior by assuming first best–response adjustment which is then perturbed by some form of noise. Instead, we consider a simple adjustment rule that connects probabilities to switch from one strategy to the other and payoffs of both strategies directly. Roughly said, the main assumptions are the following (see below): (1) agents stay with their strategy in case of successful coordination with their neighbors, (2) agents are more likely to switch if the other strategy earns relatively higher payoffs.

Technically we model individual probabilities to adjust a strategy by so–called flip rates. These rates are real–valued functions and determine the probability for an agent to switch (flip) to the other strategy within an infinitesimally short period of time. Precisely this works as follows. Denote $r^s(x, \xi_t) \in [0, \infty)$ the flip rate of agent x given the state of the population ξ_t with x himself playing strategy s. Then for $\delta \downarrow 0$ it holds that

$$Prob[\xi_{t+\delta}(x) \neq s] = r^s(x, \xi_t) \cdot \delta \ + o(\delta). \tag{2}$$

Thus, given ξ_t, for infinitesimally short periods of time the probability for agent x to adjust his strategy within that period from $s \in \{T, B\}$ to the complementary strategy s^C equals the product of the flip rate $r^s(x, \xi_t)$ times the length of the time period.

The next two assumptions give the restrictions we want to make on an agent's adjustment.

Assumption 1 (Nash equilibrium) *Flip rates are zero if and only if all agents in the neighborhood coordinate on the same strategy, i.e. agents play a Nash equilibrium. Precisely, for any* $x \in \mathbb{Z}^n, \xi_t \in X, s \in \{T, B\}$

$$r^s(x, \xi_t) = 0 \ \Leftrightarrow \ \forall y \in N(x) : \xi_t(y) = s. \tag{3}$$

Assumption 2 (Flips to better strategies are more likely) *Flip rates depend on payoff differences in a linear monotonic way. The higher the relative payoff advantage of a strategy the larger the rate to flip to this strategy. Precisely, for any* $x \in \mathbb{Z}^n, \xi_t \in X$

$$r^B(x, \xi_t) \Leftrightarrow r^T(x, \xi_t) = \lambda(\pi^T(x, \xi_t) \Leftrightarrow \pi^B(x, \xi_t)), \tag{4}$$

where $0 < \lambda < \infty$.

Assumption 1 captures the idea that individual learning forces are weak at Nash equilibria of the game. If no single neighboring opponent plays the other strategy, this other strategy is not a best–response to any of the neighbors' currently played strategy. Hence there is no reason to play that other strategy. Thus flip rates are zero. In this situation an agent's behavior coincides with pure best–response behavior.

Assumption 2 is the important behavioral assumption in our model. It is motivated by the idea that agents do not over–sensitively react to changes in their local environment by always adjusting their strategy towards best–responses with probability one. Instead, agents are assumed to follow a rule of *adjusting towards better strategies more likely*. The larger the payoff difference between the other strategy and the current strategy is the more likely it is to flip to the other strategy. The sensitivity of this relation is governed by the parameter λ, which we assume to be finite.[6] The larger λ the more sensitive the adjustment with respect to payoff differences. Since payoffs in the underlying coordination game are finite as well, flip rates are always finite. This ensures, by (2), that agents are locked in for infinitesimally short periods of time. During these time periods they are 'programmed' to the chosen strategy as mentioned above.

Note that under Assumption 2 the probability to flip is not necessarily zero whenever the payoff of the other strategy is less than the one that is currently earned. Assumption 1 says that if there is at least one neighbor who plays the other strategy there is also a strictly positive probability for the respective agent to switch to that strategy, even if it makes him worse off. Yet by Assumption 2 it follows that after he has switched to a bad strategy the probability to switch back again to the good strategy is always higher than the one before.

A feature of our approach is that with probability zero two agents flip at exactly the same time. Hence, individual strategy adjustments are non–synchronized, which allows us to ignore the effects of simultaneous strategy revision.

5. Relation to Other Work

In a static framework Rosenthal (1989) has studied an idea similar to our approach. For general two person games with finite numbers of strategies the author explores a solution concept where, instead of playing best–responses with probability one, players use "better responses with probabilities not lower than worse responses" (op.cit., p274). Using a notation similar to ours this idea can be made precise as follows. Let p_i and p_j denote the probabilities with which a player intends to use his strategies i and j. Let π_i and π_j denote the payoffs of strategies i and j given some chosen strategy of the other player. Then Rosenthal assumes that if $p_i > 0$ and $p_j > 0$

$$p_i \Leftrightarrow p_j = \lambda(\pi_i \Leftrightarrow \pi_j), \tag{5}$$

where λ is a finite parameter playing the same role as in our model. Comparing equation (5) to our Assumption 2 shows that our model can in fact be seen as a dynamic version of Rosenthal's model of boundedly rational behavior. Instead of relating static choice probabilities to payoff differences we assume that probabilities to *change* a strategy are connected to corresponding differences in payoffs.

[6]Note that $\lambda = \infty$ would correspond to a pure best–response scenario since (2) shows that a flip rate equal to infinity implies an instantaneous adjustment if the other strategy earns a larger payoff.

If we look at an agent's problem from a decision theoretic point of view, the idea of Assumption 2 is also closely related to the notion of a *proportional imitation rule* as introduced by Schlag (1998). There an imitation rule is called proportional if the difference in probabilities of switching from strategy i to strategy j and vice versa is proportional to the payoff difference between strategies i and j. In Schlag's model payoffs are determined via a multi–armed bandit and an agent can learn other strategies and payoffs by sampling other agents. He then (randomly) decides to imitate, i.e. switch to the other agent's strategy, or not. This is, of course, different to our model since agents do not imitate, nor do they observe other agents' payoffs. Still, given an agent's information about his current strategy's payoff and the payoff of another strategy (be it via sampling or by own calculation) our main assumptions coincide in the sense that probabilities to switch are proportional to the difference in payoffs between both strategies.

6. Definition of Flip Rates

In order to bring Assumptions 1 and 2 into being recall the payoff matrix of the underlying coordination game (Figure 1). Given a configuration ξ_t the payoff agent x earns from playing strategy T or B can be computed as

$$\begin{aligned} \pi^T(x,\xi_t) &= \sum_{\substack{y\in N(x)\\ \xi_t(y)=T}} a + \sum_{\substack{y\in N(x)\\ \xi_t(y)=B}} c, && (6)\\ \pi^B(x,\xi_t) &= \sum_{\substack{y\in N(x)\\ \xi_t(y)=T}} d + \sum_{\substack{y\in N(x)\\ \xi_t(y)=B}} b. && (7) \end{aligned}$$

Hence payoff differences are equal to

$$\pi^T(x,\xi_t) \Leftrightarrow \pi^B(x,\xi_t) = (a \Leftrightarrow d)N^T(x,\xi_t) \;\; \Leftrightarrow \;\; (b \Leftrightarrow c)N^B(x,\xi_t), \tag{8}$$

where $N^{\hat{s}}(x,\xi_t)$ gives the number of agent x's neighbors who play strategy $\hat{s}$ when the state of the population is determined by configuration ξ_t.

In view of Assumptions 1 and 2 there are still many possibilities to define flip rates, as they leave some degrees of freedom. The first assumption gives boundary conditions while the second one fixes relative values only. We will stick to the simplest version possible, which is based on differences as given in equation (8).

Definition 1 (Stochastic Strategy Adjustment) *For $x \in \mathbb{Z}^n$ and $\xi_t \in X$ define*

$$\begin{aligned} r^B(x,\xi_t) &= \lambda(a \Leftrightarrow d)N^T(x,\xi_t), && (9)\\ r^T(x,\xi_t) &= \lambda(b \Leftrightarrow c)N^B(x,\xi_t), && (10) \end{aligned}$$

where $0 < \lambda < \infty$ and a, b, c, d are payoffs in the coordination game.

Definition 1 is a clear aggregation of Assumptions 1 and 2. Flip rates are zero whenever all neighbors of x play the same strategy as x. For every configuration ξ_t the difference between flip rates equals the difference of payoffs times the sensitivity parameter λ.

Note that our assumptions imply that every additional neighbor that plays the other strategy increases the probability to switch to that strategy by a value equal to the equilibrium payoff of that strategy, and simultaneously decreases the probability by a value equal to the off–equilibrium payoff of that strategy. The increase can be seen as corresponding to the possibility to earn the equilibrium payoff after a switch to that strategy. The decrease of the probability is then related to a simultaneous loss of the off–equilibrium payoff that is currently earned. In this sense factors $(a - d)$ and $(b - c)$ capture revenue minus opportunity cost of adjusting from one strategy to the other. Of course, since agents play a coordination game these terms are always positive. Thus, if there is at least one neighbor playing the other strategy, the probability of switching to that strategy is positive, as well. However, the concrete likelihood of an actual switch is determined by the magnitude of these terms.

This suggests that we have to distinguish between two cases, either $(a - d) = (b - c)$ or $(a - d) > (b - c)$.[7] In the first case we are in a symmetric situation. The probability of switching depends just on the number of neighbors that play the other strategy, equally weighted for both strategies. In the language of Harsanyi and Selten (1988) this case is equivalent to saying that both equilibria are equally risky while if $(a - d) > (b - c)$ strategy profile (T, T) is the *risk–dominant* equilibrium. In our model $(a - d) > (b - c)$ implies that we are in an asymmetric situation, where strategy T is weighted more strongly, which may suggest already the direction of play within our population. We will restrict attention to this case for the rest of this paper. However, before we do so, we quickly want to mention its relation to the symmetric case.

Using common results from the theory of interacting particle systems (see Liggett (1985) for a good introduction) it can easily be shown that flip rates in Definition 1 define a unique Markov process $\{\xi_t\}_{t \geq 0}$ on the state space of all configurations X. In the following we call this process the *adjustment process*. In the symmetric case, where $(a - d) = (b - c)$, the behavior of this process is equivalent to the behavior of the so–called *voter model*, which was introduced independently by Clifford and Sudbury (1973) and Holley and Liggett (1975). In the asymmetric case, where $(a - d) > (b - c)$, the process is equivalent to the so–called *biased voter model*. This process was first considered by Schwartz (1977) and later by Bramson & Griffeath (1980, 1981). Bramson & Griffeath looked for results concerning the evolution of the process to describe the possible spread of cancerous cells, an approach that was introduced by Williams & Bjerknes (1972), while Schwartz was more interested in the duality theory of a larger class of Markov processes.

We now turn to an analysis of the evolution of play when agents adjust according to Definition 1 and $(a - d) > (b - c)$.

[7] The case $(a - d) < (b - c)$ is equivalent to the case $(a - d) > (b - c)$ by changing names of the strategies.

7. The Sensitivity Parameter λ

Since the behavior of the adjustment process is equivalent to that of the biased voter model results on the former follow from results on the latter. We therefore skip a reproduction of corresponding proofs and instead provide an intuition.

The first observation is that as long as λ is finite its value plays no role for the long run behavior ($t = \infty$) of the process.

Proposition 1 *For $\lambda < \infty$ the long run behavior of the adjustment process is independent of λ.*

Proof: The claim follows immediately from the fact that a change of λ simply results into a change of the time scale and that properties concerning long run behavior ($t = \infty$) are independent of the time scale. □

In the short and in the medium run the value of λ does, of course, play a big role for the behavior of the process. For example, the expected number of agents that play strategy T at any finite time t_0 does depend on the concrete value of λ. Low values of λ create high inertia within the adjustment of an agent, while high values speed up the evolution of play. However, since technically any effects of a change of λ correspond to a rescaling of time, a change has no qualitative implications. Convergence is the same for every finite λ.[8] Since we will focus on long run behavior of the process only we normalize $\lambda = 1$.

8. Clustering

One of the most important problems is, of course, the characterization of the set of invariant distributions of the adjustment process $\{\xi_t\}_{t\geq 0}$, since these will be the only possible limiting distributions for the process. Obviously, the prominent measures ν_B and ν_T that correspond to the strict Nash equilibria (B, B) and (T, T), concentrating on the states where everybody plays B (denoted as $\mathbf{B}$) and everybody plays T (denoted as $\mathbf{T}$) respectively, are both invariant. So the process will never be ergodic. Once the process is in one of these states it will never leave it again as they are both absorbing states.

This fact corresponds to the results of KMR (1993), Young (1993), and others before introducing mutations. With pure best–response behavior either states where the whole population plays one of the two strict Nash equilibria are absorbing states. Only after the noise component is added a selection between these states occurs. In our model, as we will see, a selection occurs already on the basis of stochastic strategy adjustment.

[8]In this sense the parameter works in a similar way as the inertia parameter μ in the learning model of Hart and Mas-Colell (1997).

Certainly also every convex combination of ν_B and ν_T is invariant as in general the set of invariant distributions is a compact convex set. So the question is, if besides ν_B and ν_T there exists any other extreme invariant distribution ν for the process. The answer is given in the following proposition.

Proposition 2 *The only extreme invariant distributions of the adjustment process $\{\xi_t\}_{t\geq 0}$ are ν_B and ν_T, that concentrate on* **B** *and* **T***, respectively.*

The key to Proposition 2 lies in the analysis of the so–called dual process.[9] For the adjustment process this process is a continuous time particle jump process on $\mathbb{Z}^n$ where each particle jumps with rate $(b \Leftrightarrow c)$ to a neighboring site and also produces a particle in an unoccupied site with a rate equal to $(a \Leftrightarrow d) \Leftrightarrow (b \Leftrightarrow c)$. If a particle attempts to occupy a site that is already occupied the two particles coalesce. The result then follows from Schwartz (1977) who has shown that whenever the dual process of a Markov process on X is monotone and fulfils a certain growth condition, then any invariant distribution must be a convex combination of the two measures ν_B and ν_T.[10]

Let $\{\xi_t^\mu\}_{t\geq 0}$ be the adjustment process that starts with initial distribution μ and let μ_t denote the distribution of that process at time t. An immediate consequence of Proposition 2 is that if $\lim_{t\to\infty}\mu_t$ exists, the process clusters, that is for any $x, y \in \mathbb{Z}^n$ the probability of $\{\xi_t^\mu(x) \neq \xi_t^\mu(y)\}$ converges to 0 as t goes to infinity. Thus, we obtain the following corollary.

Corollary 1 *The only long run configurations are those where all agents coordinate on one of the two Nash equilibria.*

Since long run configurations are homogenous the next question is, on which equilibrium agents will coordinate more likely. In other words, for which initial distributions will agents coordinate on the risk–dominant equilibrium? The answer to these questions will give the desired selection result.

9. Coordination on the Risk–Dominant Equilibrium

Note that invariant distributions of the adjustment process are translation invariant, where the latter is defined as follows.

Definition 2 *A probability measure μ on X is* translation invariant *if for any finite collection of agents $(x_1, \ldots, x_k)$, any profile of strategies $(i_1, \ldots, i_k)$, with $i_j \in \{T, B\}$, and $z \in \mathbb{Z}^n$*

$$\mu\Big(\xi(z+x_1)=i_1,\ldots,\xi(z+x_k)=i_k\Big) = \mu\Big(\xi(x_1)=i_1,\ldots,\xi(x_k)=i_k\Big), \tag{11}$$

[9] See Liggett (1985, chII) on duality for particle systems.

[10] In our setting a process is monotone if flip rates are increasing functions in the number of neighbors that play 'the other' strategy.

i.e. probabilities do not depend on z.

Since the dynamics of the adjustment process are translation invariant as well, in the sense that the assumed behavior is the same for every agent $x \in \mathbb{Z}^n$, this suggests that the property of translation invariance plays an important role in the model. The next result fully characterizes convergence of play given it starts with a translation invariant distribution.

Proposition 3 *Let the initial distribution μ be translation invariant. Then*

$$\lim_{t\to\infty} \mu_t = \alpha\nu_B + (1 \Leftrightarrow \alpha)\nu_T, \tag{12}$$

where $\alpha = \mu(\mathbf{B})$.

The result in Proposition 3 is proved in several steps. The first observation is that $\frac{d}{dt}\mu_t(\xi(x) = T) = ((a \Leftrightarrow d) \Leftrightarrow (b \Leftrightarrow c)) \sum_{y\in N(x)} \mu_t(\xi(x) = B,\, \xi(y) = T)$. Since $(a \Leftrightarrow d) > (b \Leftrightarrow c)$, the latter is non–negative, hence the probability for playing T at time t is non–decreasing in t. Since $\mu_t \in [0, 1]$, it must converge and consequently, the process clusters. The fact that $\alpha = \mu(\mathbf{B})$ follows from the translation invariance of μ and hence of μ_t, which implies that $\mu(\mathbf{B}) = \mu_t(\mathbf{B})$. See Schwartz (1977) for details.

Equation (12) nicely states the long run effects of interaction in our model. In general, any limiting distribution must be a convex combination of the two measures ν_B and ν_T. Now for a translation invariant distribution μ the parameter α that determines the mixture between these measures is already uniquely determined by the value $\mu(\mathbf{B})$, which is the probability that *all* agents initially start with playing strategy B. Once this probability is zero we obtain convergence to ν_T. On the other hand, convergence to ν_B is obtained only in the case where the process starts already in that particular state, i.e. $\mu(\mathbf{B}) = 1$. This is a quite substantial result which is reformulated in the following corollary.

Corollary 2 *If the initial distribution of the process is translation invariant, and almost surely at least one agent plays strategy T at the beginning, then with probability one agents coordinate on the risk–dominant equilibrium.*

As an example for the last result consider the process that starts with initial distribution μ_θ, being the Bernoulli product measure with parameter θ where for each agent $x \in \mathbb{Z}^n$, $\mu_\theta(\xi(x) = T) = \theta$. Certainly μ_θ is translation invariant. If $\theta > 0$, $\mu_\theta(\mathbf{B}) = 0$, hence agents will coordinate on the risk–dominant equilibrium. While the state of the population at the beginning is characterized by individual independence, in the long run the evolution of the process in time eventually leads to complete unanimity. More than that, all players eventually agree to coordinate on the risk–dominant equilibrium. The driving force that makes this coordination possible is the adjustment mechanism determined by the interaction between players. Even though this interaction is locally restricted to the neighborhood of a player, because of the considerable overlap between these neighborhoods its effect is on the population as a whole.

Remark: The above result is an immediate implication of $(a - d) > (b - c)$, i.e. strategy T being risk–dominant. Denote $\kappa = \frac{a-d}{b-c}$ the ratio of these terms, measuring the degree of risk–dominance. An important effect of κ approaching 1, i.e. both equilibria becoming equally risky, is that the expected waiting time for the process to hit *any* absorbing state, either **T** or **B**, can take very large values. This result is due to Cox (1989) and holds for finite populations where, contrary to the infinite case, the probability to hit an absorbing state in finite time equals 1. Consider, for example, a finite population of N agents located on the torus imbedded in the 2–dimensional lattice $\mathbb{Z}^2$. Let the initial distribution of play be given by the finite version of μ_θ with $\theta > 0$. Then, as κ approaches 1, the expected waiting time to hit *any* equilibrium state tends to

$$\frac{2}{\pi} N^2 \log N(-\theta \log \theta - (1-\theta)\log(1-\theta)) \tag{13}$$

as N becomes large.[11] See Cox (1989) for more. Thus, when both equilibria are equally risky and the population is large stochastic strategy adjustment needs very long waiting times until the whole population coordinates on either of the Nash–equilibria, even though interaction is restricted to local neighborhoods.

10. The Spread of Risk–Dominant Play

So far results have been obtained for the case when the initial state of the population may be described by translation invariant distributions. It is interesting to see how risk–dominant play spreads starting from an arbitrary set of agents that play strategy T at time zero. Here lies a great advantage of any spatial model over those like KMR (1993) and Young (1993), where the population is not endowed with a spatial structure and therefore the possibility to study a real spread of a strategy is not given. Denote the set of agents that play T a time zero by $A \subset \mathbb{Z}^n$. Obviously, a spread of risk–dominant play can occur only if the absorbing state **B**, where all agents play B, is never reached. Let ξ_t^A be the adjustment process that starts with initial distribution δ_A, with $A \subset \mathbb{Z}^n$, and let τ^A denote the hitting time of the state **B** for the process ξ_t^A. The next propositions calculates the probability for the process ξ_t^A to reach this state in finite time.

Proposition 4 *Let $A \subset \mathbb{Z}^n$. Then*

$$Prob[\tau^A < \infty] = \begin{cases} \kappa^{-|A|} & \textit{if} \quad A \ \textit{finite} \\ 0 & \textit{if} \quad A \ \textit{infinite}, \end{cases} \tag{14}$$

where $\kappa = \frac{a-d}{b-c}$.

[11] The order of limits in this statement is first $\kappa \downarrow 1$ and then $N \uparrow \infty$.

Since we are not aware of any proof of Proposition 4 we shall give a precise proof in the appendix. The proposition shows that the probability to hit the absorbing state $\mathbf{B}$ in finite time depends on the number of agents that play T at the beginning and the ratio of individual weights of adjustment $\kappa = \frac{a-d}{b-d}$. Since T is risk–dominant this ratio is always larger than 1. So the probability decreases exponentially as the number of initial T–strategists grows. Note that the probability does not depend on the spatial spread of A, i.e. how densely these agents are actually distributed within the population. The only thing that matters is the cardinality of A. In particular, if A is infinite we obtain again almost sure coordination on the risk–dominant equilibrium.

In (14) the degree of risk–dominance, expressed by κ, directly enters the equation. The more risk–dominant strategy T is, the faster the probability to enter the state $\mathbf{B}$ decreases as the size of A grows. In the other direction, as both equilibria become equally risky (κ approaches 1) the probability to reach the equilibrium where everybody plays B tends to 1 for finite A.

Now suppose that strategy T is 'very' risk–dominant in the sense that $\kappa >> 1$. By Proposition 4, for large A the event $\{\tau^A = \infty\}$ has overwhelming probability. Denote $D_r = \{x \in \mathbb{Z}^n : |x| \leq r\}$ the ball of radius r around the origin and let $\Lambda(\xi_t^A) \subset \mathbb{Z}^n$ be the set of agents that play strategy T given configuration ξ_t^A. The next proposition shows that, conditioned on $\{\tau^A = \infty\}$, strategy T eventually spreads at least linearly. Thus again, if A is infinite risk–dominant play almost surely overtakes the whole population in a linear fashion.

Proposition 5 *For every set $A \neq \emptyset$ of agents playing strategy T at time zero,*

$$Prob[\exists t_0 < \infty \ \forall t \geq t_0 \ D_{\beta t} \subset \Lambda(\xi_t^A) |\ \tau^A = \infty] = 1. \tag{15}$$

For a proof of Proposition 5 see Bramson and Griffeath (1981), who also show that the constant β depends only the dimension n and the parameter κ. It is instructive to compare the result to a similar one in the equilibrium selection model of Hofbauer (1998). There, the author studies a travelling wave approach to define a *spatially dominant* equilibrium that is shown to coincide with risk–dominance in symmetric 2×2 coordination games. A notable observation is that the speed of his wave is closely related to the asymptotic growth of the set of T–players in our model. Consider, for example the simple coordination game with off–equilibrium payoffs c and d equal to zero. Then it can be shown that the asymptotic speed of a spread of T under the adjustment process is equal to $\frac{1}{4n\sqrt{n}} \frac{a-b}{a+b}$ (Bramson and Griffeath, 1981). The wave speed in Hofbauer (1998) using replicator dynamics as reaction dynamics equals $\sqrt{\frac{e}{2}} \frac{a-b}{\sqrt{a+b}}$, where e captures an additional migration rate of agents. In Hofbauer's model the population is distributed on the one-dimensional continuum $\mathbb{R}$. Hence, taking the same dimension $n = 1$ in our model, risk–dominant play spreads at a similar speed as in the model of Hofbauer. At the same time the underlying geometric structure of the population is, of course, substantially different. This suggests that the result in Proposition 5 does not depend on the special structure of the space $\mathbb{Z}^n$ as one might, perhaps, have suspected.

11. Conclusion

We have studied a new form of strategy adjustment behavior by agents who repeatedly play a symmetric 2×2 coordination game with local neighbors. Rather than considering perturbations of pure best–response we focus on the original idea of best–response, which says that agents are influenced by payoff differences. We smooth this influence by assuming that, instead of switching to best responses with probability one, agents switch to better responses more likely. The underlying idea of this assumption corresponds to the bounded rationality model of Rosenthal (1989) and the notion of proportional imitation of Schlag (1998). Based on a spatial model of local interaction our results say that agents are in fact more likely to coordinate on the risk–dominant equilibrium. In this sense our approach supports the results of KMR (1993), Young (1993), and others. Precisely, in our model risk–dominant play prevails with probability one whenever the initial fraction of agents that play the risk–dominant strategy has at least some positive measure. Since our population is located on the infinite lattice $\mathbb{Z}^n$, this holds whenever the initial fraction contains infinitely many agents, independent of the spatial distribution of these agents. This is obtained, e.g., by starting either with a corresponding Dirac measure or with independent probability assignments to each player, where strategy T has at least some positive probability.

Our model shares with others the special feature that agents do not have the possibility to influence the set of opponents they face. Every agent interacts with a local neighborhood that is exogenously given and remains fixed forever. In contrast, Ely (1995) and Bhaskar and Vega-Redondo (1997) have shown that once agents are allowed to choose their set of opponents the situation looks totally different. They introduce a number of available locations where agents can meet and exclusively play the game with each other. Thus the choice of a location directly determines the set of opponents an agent is going to face. The effect is that agents will no longer coordinate on the risk–dominant but instead choose the payoff–dominant equilibrium.[12]

This suggest that results on equilibrium selection do not only depend upon the characteristics of the noise process or the considered adjustment behavior of an agent but also, and perhaps even more, on the specific kind of interaction that is assumed. In consequence, the next questions are: once interaction structures are modelled endogenously, how do these structures evolve themselves? What are the mechanisms that exist between playing specific strategies and interacting with specific neighbors? How do these mechanisms work? Do strategies perhaps arise as a direct consequence of interaction patterns? Or, in other words, do interaction patterns define the (local) environment in such a way that specific play can be observed that would not be observed if the interaction patterns were different? Promising research in this direction is already started by Morris (1996), Berninghaus and Ehrhart (1998), and Young (1998).

[12] Bhaskar and Vega-Redondo (1997) obtain the payoff–dominant equilibrium for stag-hunt games and both equilibria for pure coordination games. See our footnote 4.

Appendix

Proof of Proposition 4: Consider the case when A is finite. If t_m denotes the time of the m'th flip of ξ_t^A, then $R_m^A = |\{x \in \mathbb{Z}^n | \xi_{t_m}^A(x) = T\}|$ is a random walk on $\mathbb{N}$ where

$$R_m^A \Leftrightarrow R_{m-1}^A = \begin{cases} 1 & \text{with prob.} \quad p \\ \Leftrightarrow 1 & \text{with prob.} \quad 1 \Leftrightarrow p, \end{cases}$$

and probabilities are $p = \frac{\kappa}{\kappa+1}$ and $1 \Leftrightarrow p = \frac{1}{\kappa+1}$. This follows from the flip rates of the adjustment process: $\frac{p}{1-p} = \frac{a-d}{b-c} = \kappa$.

For $b \in \mathbb{N}$ define the stopping time $S = \min_{m \geq 0}\{R_m^A \notin (0,b)\}$, i.e. the stopping time of the first moment the random walk leaves the interval $(0,b)$. By the lemma of Borel-Cantelli it follows that S is almost surely finite. Because of the positive drift $\frac{\kappa-1}{\kappa+1}$ towards infinity, R_m^A is not a martingale. Therefore define the function $h(x) = (\frac{1-p}{p})^x$. $h(x)$ is a harmonic function with respect to R_m^A, i.e. $(Kh)(x) = h(x)$, where $K(x,\cdot)$ denotes the transition function of the random walk R_m^A. Now $H_m^A = h(R_m^A)$ is in fact a martingale and since $H_{m\wedge S}$ is clearly bounded we may use the martingale stopping theorem to follow that $E[H_S] = E[H_0]$. With p_0 and p_b, denoting the probability for R_S^A to be 0 and b, respectively, this is equivalent to

$$\begin{array}{rrcl} & p_0 h(0) + p_b h(b) & = & h(|A|) \\ \iff & p_0(h(0) \Leftrightarrow h(b)) & = & h(|A|) \Leftrightarrow h(b) \\ \iff & p_0 & = & \dfrac{h(|A|) \Leftrightarrow h(b)}{h(0) \Leftrightarrow h(b)} \\ \iff & p_0 & = & \dfrac{(\frac{1-p}{p})^{|A|} \Leftrightarrow (\frac{1-p}{p})^b}{1 \Leftrightarrow (\frac{1-p}{p})^b}. \end{array}$$

Now let the interval $(0,b)$ grow to infinity, then $P[\tau^A < \infty] = \lim_{b\to\infty} p_0$. Since the latter limes is equal to $(\frac{1-p}{p})^{|A|} = \kappa^{-|A|}$, this proves the finite case.

The infinite case now follows from the finite one by exhausting the space $\mathbb{Z}^n$ via a sequence of finite boxes with increasing radius centred around the origin, using monotonicity of the process. $\square$

References

Bergin, J. and Lipman, B.L. (1996) "Evolution with State-Dependent Mutations," *Econometrica*, 64, 943-956.

Berninghaus, S.K. and Ehrhart K.-M. (1998) "Long–run Evolution of Local Interaction Structures in Games," mimeo, Institute of Statistics and Mathematical Economics, University of Karlsruhe.

Bhaskar, V. and Vega-Redondo, F. (1997) "Migration and the Evolution of Conventions," mimeo, Department of Economics, University of Alicante.

Blume, L.E. (1993) "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5, 387-424.

Blume, L.E. (1994) "How Noise Matters," mimeo, Department of Economics, Cornell University.

Blume, L.E. (1995) "The Statistical Mechanics of Best-Response Strategy Revision," *Games and Economic Behavior*, 11, 111-145.

Bramson, M. and Griffeath, D. (1980) "On the Williams-Bjerknes Tumor Growth Model II," *Mathematical Proceedings of the Cambridge Philosophical Society*, 88, 339-357.

Bramson, M. and Griffeath, D. (1981) "On the Williams-Bjerknes Tumor Growth Model I," *Annals of Probability*, 9, 173-185.

Clifford, P. and Sudbury, A. (1973) "A Model for Spatial Conflict," *Biometrika*, 60, 581-588.

Cox, J.T. (1989) "Coalescing Random Walks and Voter Model Consensus Times on the Torus in $\mathbb{Z}^d$," *Annals of Probability*, 17, 1333-1366.

van Damme, E. and Weibull, J.W. (1998) "Evolution with Mutations Driven by Control Costs," Working Paper No. 501, Research Institute of Industrial Economics, Stockholm.

Ellison, G. (1993) "Learning, Local Interaction, and Coordination," *Econometrica*, 61, 1047-1071.

Ely, J. (1995) "Local Conventions," mimeo, Economics Department, University of California at Berkeley.

Harsanyi, J.C. and Selten, R. (1988) *A General Theory of Equilibrium Selection in Games*, MIT Press.

HART, S. AND MAS-COLELL, A. (1997) "A Simple Adaptive Procedure Leading to Correlated Equilibrium," Discussion Paper No. 126, Center for Rationality and Interactive Decision Theory, Hebrew University of Jerusalem.

HOFBAUER, J. (1998) "The Spatially Dominant Equilibrium of a Game," mimeo, Institute of Mathematics, University of Vienna.

HOLLEY, R. AND LIGGETT, T.M. (1975) "Ergodic Theorems for Weakly Interacting Systems and the Voter Model," *Annals of Probability*, 3, 643-663.

KANDORI, M., MAILATH, G.J. AND ROB, R. (1993) "Learning, Mutation and Long Run Equilibria in Games," *Econometrica*, 61, 29-56.

KIM, Y.-G. AND SOBEL, J. (1995) "An Evolutionary Approach to Pre-Play Communication," *Econometrica*, 63, 1181-1193.

LIGGETT, T.M. (1985) *Interacting Particle Systems*, Springer-Verlag New York.

MORRIS, S. (1996) "Contagion," mimeo, Department of Economics, University of Pennsylvania.

ROSENTHAL R.W. (1989) "A Bounded-Rationality Approach to the Study of Noncooperative Games," *International Journal of Game Theory*, 18, 273-292.

SCHLAG, K.H. (1998) "Why Imitate, and If So. How? A Boundedly Rational Approach to Multi-armed Bandits," *Journal of Economic Theory*, 78, 130-156.

SCHWARTZ, D.L. (1977) "Applications of Duality to a Class of Markov Processes," *Annals of Probability*, 5, 522-532.

WILLIAMS, T. AND BJERKNES, R. (1972) "Stochastic model for abnormal clone spread through epithelial basal layer," *Nature*, 236, 19-21.

YOUNG, H.P. (1993) "The Evolution of Conventions," *Econometrica*, 61, 57-84.

YOUNG, H.P (1998) "Diffusion in Social Networks," mimeo, Department of Economics, Johns Hopkins University, Baltimore.